J2
409/2:67
c.2

CHES STUDIES

A SPECIAL REPORT SERIES BY THE N.C. DEPARTMENT OF ENVIRONMENT, HEALTH, AND NATURAL RESOURCES
STATE CENTER FOR HEALTH AND ENVIRONMENTAL STATISTICS
P.O. BOX 29538, RALEIGH, N.C. 27626-0538

No. 67 June 1992

THE QUALITY OF CANCER INCIDENCE DATA IN NORTH CAROLINA

An Evaluation of the Central Cancer Registry 1990 Database

by

Carol R. Dickinson and Tim E. Aldrich

ABSTRACT

The North Carolina Central Cancer Registry (CCR) has completed its first year of statewide data collection. During 1990, 127 of 132 facilities submitted cancer incidence data; 29 of those were tumor registries. For the facilities without tumor registries, the CCR conducted a complete review of all incidence reports for quality assurance and completeness, and provided training to all new staff members. This quality control process led to two-thirds of these facilities having an accuracy rate in excess of 96 percent. For facilities with new tumor registrars, CCR offered training sessions and a 100 percent evaluation of their abstracts for at least two months.

The CCR conducted a study to assess the reliability of the data submitted. From the 127 reporting facilities, 87 persons participated in an abstracting exercise of four sample charts. Tumor registrars were consistently more accurate than the data collectors from hospitals without tumor registrars. Items presenting the most problems for all persons were the determination of the exact anatomic primary site and summary staging. Medical records personnel who abstract cancer data in hospitals without tumor registries also need to improve on the accuracy of the reporting of ethnicity, date of diagnosis, date first seen at their facility, and the sequence coding for multiple primary tumors.

In addition to the above quality control activities, the CCR performed two evaluations of the completeness of case ascertainment. First, a standard measure of completeness of reporting, the incidence/mortality ratio, was calculated for all cancers by race-sex group. This measure consistently agreed with national standards and those from the state of Georgia. Also, a casefinding audit was performed for a sample of hospitals. The results highlighted areas that need strengthening in the future.

This report is presented as documentation of the quality control activities for the CCR 1990 incidence database and the results of those activities.

INTRODUCTION

The North Carolina Central Cancer Registry (CCR) was established in 1986 and collects basic demographic and diagnostic information on all newly diagnosed cancer cases throughout the state. A multi-phase system was developed to implement statewide reporting, and from 1987 through 1989, the 18 hospitals with the larger tumor registries reported to the CCR. At the end of that period, just under 50 percent of the state's estimated cancer cases were being collected.

In 1990, the Central Cancer Registry (CCR) moved into a more established phase of reporting, and facilities without tumor registries began sending incidence reports. In June 1990 the Quality Control Unit was established and a systematic program was developed to ensure accurate, timely reporting of incidence records in line with national standards (Hilsenbeck 1990). This report provides an overview of the quality control procedures used by the CCR and the results from those procedures.

METHODS OF CANCER INCIDENCE DATA COLLECTION AND QUALITY CONTROL

Hospitals that report cancer incidence data to the CCR are classified into two categories: incidence reporting hospitals and tumor registry hospitals. This section provides a description of the quality control procedures used by the CCR for each of these groups.

Incidence Reporting Hospitals

Using the medical record as the primary data source, incidence reports are completed by hospital personnel and are submitted to the CCR with a copy of the pathology report. When received, the submission is recorded in a log book and the primary site and morphology sections are coded using the International Classification of Disease for Oncology-Field Trial Edition, 1988. Reports are visually reviewed for completeness and accuracy. A summary of error items for each hospital is forwarded to the CCR field staff representative who works with the hospital, and hospital-specific quarterly error rates are calculated, to evaluate the overall quality of the data. The purpose of this activity is to identify consistent reporting errors and plan workshops and individual teaching sessions to resolve reporting problems. Site visits are initiated by the field staff representative for all hospitals with an accuracy rate of less than 98 percent, while minor reporting problems are usually corrected by phone or letter.

The 24 data items on the incidence form are inspected and the accuracy rate is calculated as follows:

$$\text{Accuracy Rate} = \frac{\text{number of errors}}{\text{number of record X 24 elements}} \text{ X } 100$$

Cancer cases reported from these incidence reporting hospitals are entered into the CCR database and passed through a series of programmed edit checks. Corrections are made as needed and the records are merged with the CCR cumulative file. A series of frequency reports are run on each variable, and missing data items are identified and corrections made.

Ninety-eight of 103 incidence reporting hospitals submitted a total of 8,431 cancer cases to the CCR in 1990 (Figure 1). These 98 hospitals include 95 whose data submissions were completed at the close of the reporting period (March 31). Three hospitals did not complete 1990 incidence reporting on schedule. Five hospitals did not report any 1990 incidence cases; four of these indicated that they did not admit any new cancer cases during the calendar year, and one hospital refused to submit data or to permit CCR staff to collect the data. This hospital has now agreed to permit CCR staff to collect their cancer incidence data beginning with 1991 data.

Of the 98 institutions that reported cancer data on the incidence report forms (i.e., hospitals without tumor registries), 64.1 percent met an acceptable level of accuracy (98% accuracy or greater). The distribution of accuracy levels is presented in Figure 2. The reader

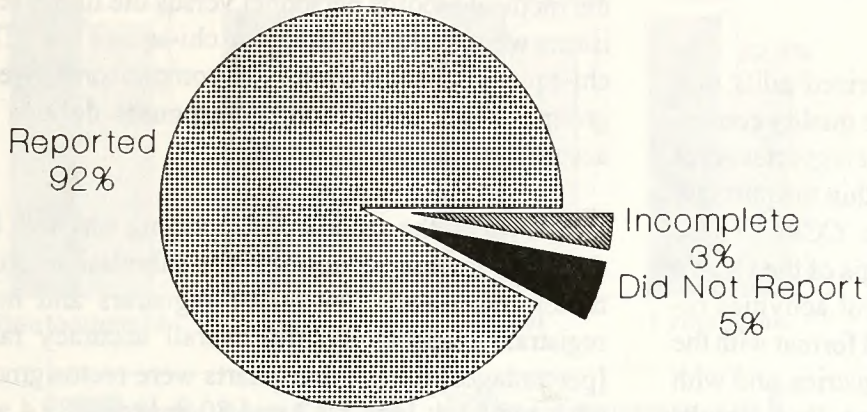

Figure 1 - Percentage Distribution of Incidence Reporting Hospitals by Reporting Status for 1990

Throughout North Carolina, there are 29 hospital-based tumor registries, an increase of 62 percent since 1986. Thirteen of these hospitals have a cancer program approved by the American College of Surgeons. In 1990, the CCR received 21,487 cases from hospital-based registries, yielding a total of 29,918 cases in the 1990 database. Data from hospital-based tumor registries are sent to the CCR primarily on diskette or magnetic tape. The data must pass a series of programmed edit checks with errors being reported back to the hospitals for correction.

should note that 17 incidence reporting facilities reporting 10 or fewer cases are not included in Figure 2. Consistent reporting errors, specifically those involving staging and identification of multiple primary tumors, as well as incomplete reports were discussed during four workshops held during 1991 for incidence reporting hospitals.

The CCR quality control program provides teaching site visits and instruction in abstracting, coding, and staging principles for tumor registrars with less than two years experience. The Tumor Registrars Association of North Carolina (TRANC) conducts educational meetings in the spring and fall of each year and operates a mentor program

Hospital-based Tumor Registries

Tumor registrars are trained oncology data specialists who maintain an extensive database including the medical history, diagnostic information, staging, treatment and life-long follow-up for all cancer patients seen at their facility. The National Tumor Registrars Association administers a certification examination to registrars (Certified Tumor Registrar - CTR) twice annually. This examination is a pivotal recognition step for the professional registrar.

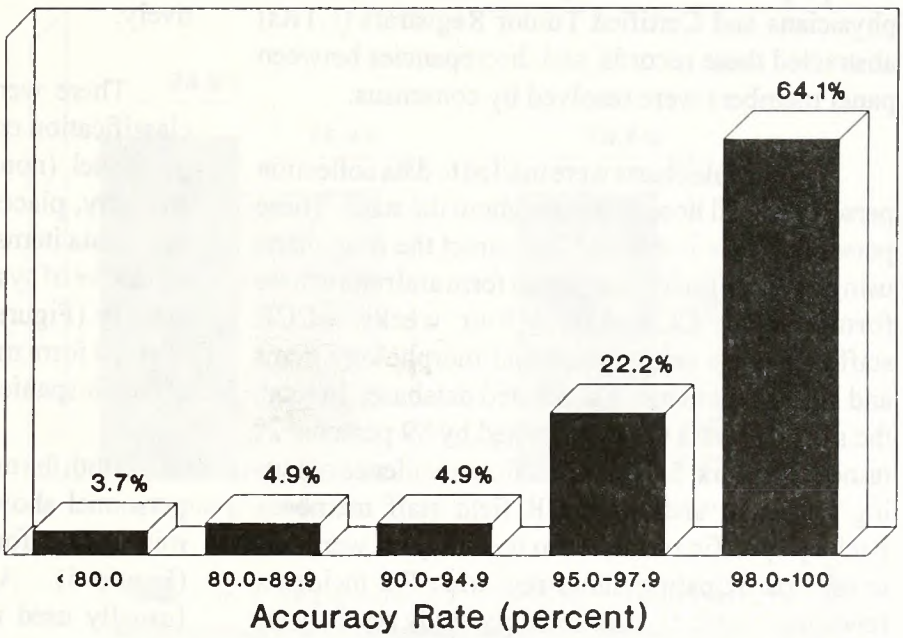

Figure 2 - Percentage Distribution of Incidence Reporting Hospitals by Accuracy Rate: Hospitals Reporting More Than 10 Cases

available to data collection personnel. The reliability audit and casefinding study described in this report are conventional quality control activities for tumor registries (Hilsenbeck 1990).

The experience with computerized edits that also may be regarded as a part of the quality control process for data submitted from tumor registries is not included in this report, as it is not within the purview of the Quality Control Unit of the CCR. These computerized edits are under the aegis of the CCR's Statistical Unit. The quality control activities reported here are consistent in style and format with the precedent of other central cancer registries, and with the Surveillance, Epidemiology and End Results (SEER) program of the National Cancer Institute (NCI) (Hilsenbeck 1990; SEER 1991).

RELIABILITY STUDY

In 1991, the Quality Control Unit undertook a study to assess the skills of the data collection personnel supplying data to the CCR, both tumor registrars and medical records personnel of the incidence reporting hospitals alike. Four standardized test charts were prepared from actual patient charts. A panel of physicians and Certified Tumor Registrars (CTRs) abstracted these records, and discrepancies between panel members were resolved by consensus.

The sample charts were mailed to data collection personnel in all hospitals throughout the state. These personnel were instructed to abstract the four charts using the CCR Incidence Report form and return these forms to the CCR within four weeks. CCR staff coded the primary site and morphology items and entered them into a simulated database. In total, the sample charts were abstracted by 89 persons: 29 tumor registrars; 58 personnel from incidence reporting hospitals; and two CCR field staff members. Facility-specific results from the analyses were sent to each participating tumor registrar. For incidence reporting hospitals, the reliability study results were used by the CCR field staff during individual site visits.

Results of the Reliability Study

Differences between the coding performance of the medical records personnel versus the tumor registrars were evaluated using the chi-square test. The chi-square statistic measures the comparison between groups as the proportion of responses defined as accurate by the consensus panel.

Overall, the accuracy of reporting was high for the demographic section and for submissions from tumor registrars. For tumor registrars and non-registrars respectively, the overall accuracy rates [percentages] for the four charts were rectosigmoid 93.1 and 83.4; lung 87.3 and 80.8; breast 89.4 and 78.9; and lymphoma 84.9 and 62.5 (Figure 3). This pattern of errors is consistent with the difficulty of the cases. The sample chart for cancer of the lung contains information that is expressly subtle regarding this complex anatomic site. Lymphomas are generally regarded as very complex cases to abstract because of the fluid nature of the tumor and the intricate staging considerations applied to this site. The item-specific errors are presented for each participant group (tumor registrars and medical records personnel of incidence reporting hospitals) and for each of the four sample charts in Tables 1-4 respectively.

There were three focal areas of demographic classification errors. Generally the medical records personnel (non-registrars) were often in error on ethnicity, place of birth and sequence number. Of these data items, errors with ethnicity were the most reflective of systematic misclassification or failure to classify (Figure 4). The order of the codes on the abstract form may have contributed to this error (i.e., not-of-hispanic-origin is code #2).

Both the tumor registrars and the medical records personnel showed considerable variation in determining the fourth digit of the primary site code (Figure 5). At the major three-digit rubric level [usually used in research analyses], accuracy was

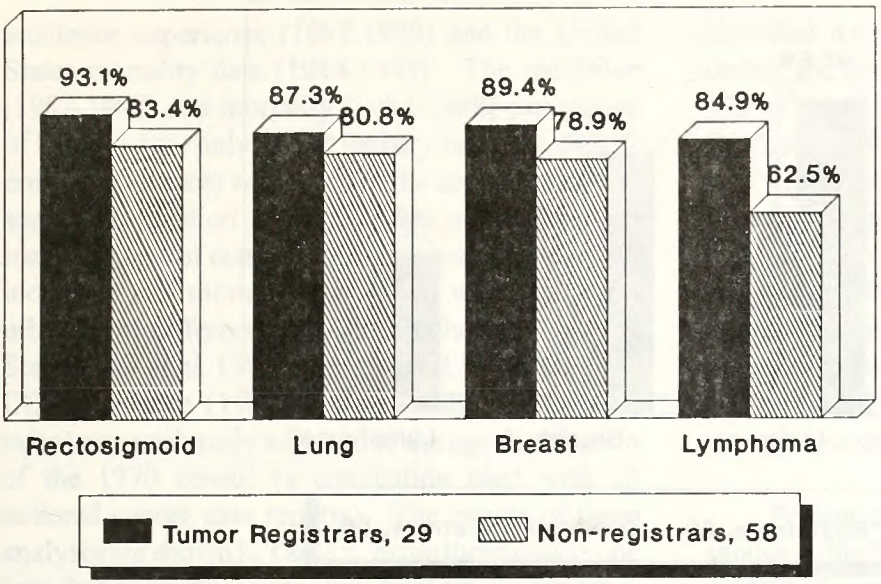

Figure 3 - Overall Accuracy Rates by Primary Site and Type of Personnel (For the Four Sample Charts Abstracted)

consistently over 98 percent. Primary site determination of the rectosigmoid case created the most difficulty for the medical records personnel; generally they coded the case to colon. The study also showed there was consistent and substantive disparity involving summary staging (Figure 6). The greatest staging error rate was for the lung cancer case. Summary staging is a subject that needs substantial attention at future training sessions. The focus of the spring 1992 meeting of the Tumor Registrars Association of North Carolina will be lung cancer.

INCIDENCE/MORTALITY RATIO

The National Cancer Institute (NCI) operates nine population-based cancer registries under its Surveillance, Epidemiology and End Results (SEER) program (SEER 1991). The CCR works closely with these SEER registries and those of other states in the American Association of Central Cancer Registries (AACCR). The sample patient charts used in the reliability audit described above were obtained from the quality control activities of the NCI programs. One measure that NCI employs as a criterion for completeness of incidence ascertainment is termed the "incidence/mortality" ratio. This value is obtained by calculating age-adjusted incidence and mortality rates for the same population and dividing one by the other. Completeness of ascertainment for mortality is assumed to be relatively constant. Variability with incidence ascertainment may then logically be indicated (within reasonable limits) by the ratio, assuming that cancer survival is relatively constant nationwide.

To assist with interpreting the preliminary North Carolina incidence/mortality ratio, two comparisons were made by race-sex group. One comparison was to the United States, represented by the combined SEER

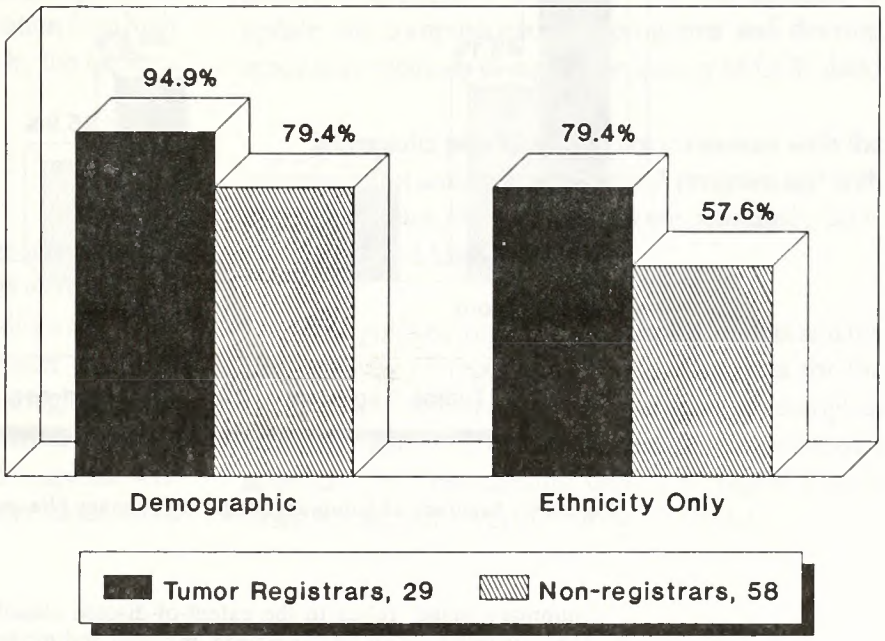

Figure 4 - Acuracy of Demographic Items by Type of Personnel

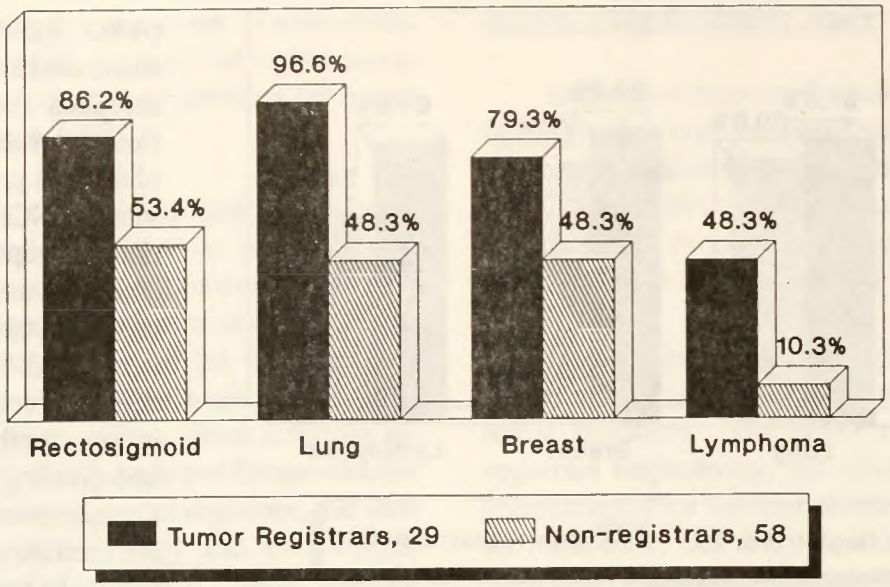

Figure 5 - Accuracy of Primary Site Code to the 4th Digit by Type of Personnel

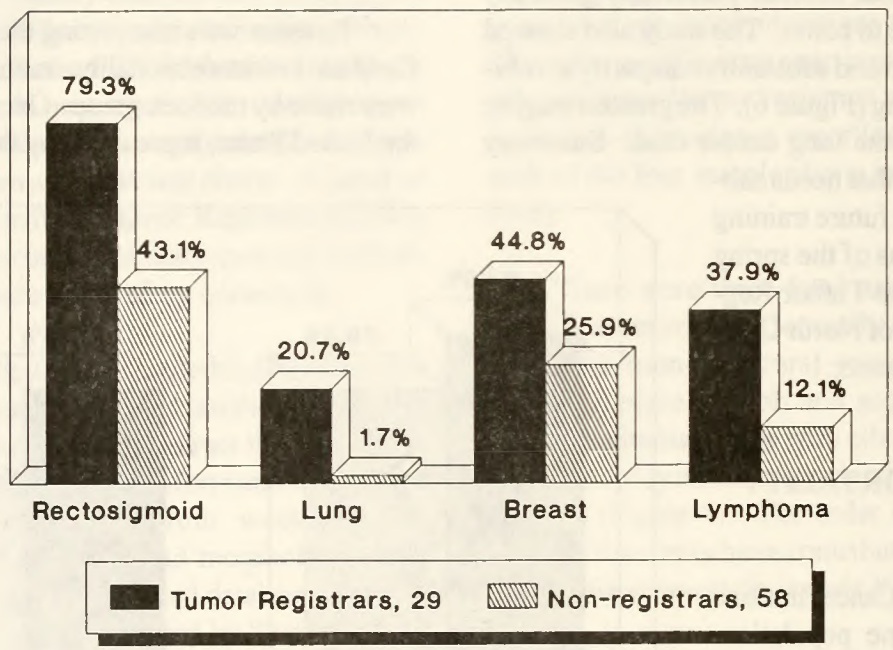

Figure 6 - Accuracy of Summary Stage* by Primary Site and Type of Personnel

*"Summary Stage" refers to the extent-of-disease classification widely used in the nation. The categories are *in situ*, localized, regional and distant. The terms are applied to describe the degree of invasion of the tumor.

incidence experience (1987-1988) and the United States mortality data (1984-1988). The incidence (1987-1988) and mortality (1984-1988) experience of Georgia (the only SEER registry in the southeastern United States) was used for the second comparison. The addition of the Georgia experience was made because of concerns for regional differences in incidence and mortality associated with racial and urban/rural differences for the southeastern United States (Liff *et al.* 1991; Monroe *et al.* In Press). The CCR incidence (1990) and mortality (1984-1988) rates were uniformly adjusted to the age distribution of the 1970 census (a convention used with all national cancer data reports). The results of these analyses are shown in Table 5. From these results one sees that individual rates vary (incidence or mortality), but all of the incidence/mortality ratios, including North Carolina's, are quite consistently in the range of 1.5 to 2.5.

The North Carolina incidence/mortality ratios for white males (2.03) and nonwhite females (2.01) are between those for SEER and Georgia, suggesting "middle of the road" completeness in incidence case ascertainment. Nonwhite males may represent the only instance for potential under-ascertainment of cases. All in all, these analyses as well as comparisons to Illinois and Utah support an interpretation for a high degree of ascertainment of cancer cases by the CCR.

CASEFINDING AUDIT

An audit was performed of the case identification in a sample of hospitals. Twelve hospitals were audited; these were chosen to represent large, medium and small institutions (in terms of numbers of cancer cases reported) and randomly for three geographic regions of the state (west, central and east). Breast cancer was the anatomic site of study. From these 12 hospitals, 412 pathology reports for 1990 breast cancer cases were identified through the audit activities of the quality control and field staff personnel of the CCR. A total of 373 of these cases had been reported to the CCR, representing 90.5 percent ascertainment. Of the remaining 39 cases, four were consultations for other hospitals who did not report the cases. Hospitals participating in the audit will receive the results and a list of the cases not reported to the CCR. This will assist them in reviewing their casefinding procedures. To assist incidence reporting facilities, methods of casefinding were emphasized and discussed by the quality control and field staff through a series of workshops in the fall of 1991.

Percent casefinding by death certificate only is another of the SEER criteria for quality assurance. This method of case ascertainment is difficult to evaluate for a single year of data (the situation in North Carolina) since mortality lags behind incidence by many years for most cancers and the only year for which complete incidence data are available is 1990.

CONCLUSION

The Central Cancer Registry has received tremendous support from the tumor registrars and hospital personnel throughout the state. We will continue to update our computerized edit programs and develop innovative methods to ensure the quality of CCR data.

The results provided here are consistent with the quality control analyses of the SEER program and with three other statewide registries that we consulted (Georgia, Illinois and Utah).

The goals of improving the quality of data and the completeness of reporting are on-going ones for the CCR. One objective toward the goal of complete ascertainment for 1992 will be the development of a schedule for free-standing clinics to report cancer cases not seen in North Carolina hospitals.

REFERENCES

Boring, C.C., Squires, T.S. and Tong. T. (1991). Cancer Statistics, 1991. *CA - A Cancer Journal for Clinicians,* 41(1):19-36.

Hilsenbeck S.G. (1990). Quality Control Practices in Centralized Tumor Registries in North America. *Journal of Clinical Epidemiology.* 43(11):1201-12.

Liff J.M., Chow W.H., Greenberg R.S. (1991). Rural-Urban Differences in Stage at Diagnosis: Possible Relationship to Cancer Screening. *Cancer* 67:1454-59.

Monroe, A.C., Aldrich T.E., Ricketts T.C., and Cooper, M.R. (In Press). Rural-Urban Differences in Cancer Stage at Diagnosis in North Carolina. *Cancer.*

SEER (1991). Cancer Statistics Review: 1973-1988. National Cancer Institute, Division of Cancer Prevention and Control, Surveillance, Epidemiology and End Results [SEER] Program, NIH Pub. 91-2789; Bethesda, MD.

TABLE 1: NUMBER AND PERCENT OF CORRECT RESPONSES BY TYPE OF RESPONDENT AND DATA ITEM FOR MEDICAL CHART 101010 - RECTOSIGMOID COLON

CORRECT RESPONSES

DATA ITEM	TUMOR REGISTRARS		MEDICAL RECORDS PERSONNEL	
	NUMBER	PERCENT	NUMBER	PERCENT
Zip Code	29	100.0	58	100.0
County Residence	28	96.6	56	96.6
Date of Birth	29	100.0	58	100.0
Race	29	100.0	58	100.0
Ethnicity	23	79.3	25	43.1**
Sex	29	100.0	58	100.0
Soc. Sec. No.	29	100.0	58	100.0
Place of Birth	25	86.2	34	58.6*
Dx Elsewhere	28	96.6	55	94.8
Sequence #	29	100.0	58	100.0
Primary Site	25	86.2	31	53.4**
PrimX (3 digits)	28	96.6	35	60.3**
Laterality	29	100.0	57	98.3
Cell Type	29	100.0	55	94.8
Grade	28	96.6	57	98.3
Confirmation	29	100.0	57	98.3
Date of Dx	23	79.3	30	51.7*
Stage at Dx	23	79.3	25	43.1**
Date 1st Seen	25	86.2	53	91.4
Stage 1st Seen	22	75.9	22	37.9**

NOTE: Statistical significance of the tumor registrar performance versus the medical records personnel performance, using a one-degree-of-freedom chi-square test: * $p < 0.05$ ** $p < 0.01$.

TABLE 2: NUMBER AND PERCENT OF CORRECT RESPONSES BY TYPE OF RESPONDENT AND DATA ITEM FOR MEDICAL CHART 202020 - LUNG

CORRECT RESPONSES

DATA ITEM	TUMOR REGISTRARS		MEDICAL RECORDS PERSONNEL	
	NUMBER	PERCENT	NUMBER	PERCENT
Zip Code	29	100.0	57	98.3
County Residence	29	100.0	57	98.3
Date of Birth	29	100.0	58	100.0
Race	28	96.6	58	100.0
Ethnicity	25	86.2	26	44.8**
Sex	28	96.6	58	100.0
Soc. Sec. No.	29	100.0	58	100.0
Place of Birth	27	93.1	44	75.9
Dx Elsewhere	25	86.2	45	77.6
Sequence #	29	100.0	58	100.0
Primary Site	28	96.6	28	48.3**
PrimX (3 digits)	29	100.0	57	98.3
Laterality	29	100.0	55	94.8
Cell Type	29	100.0	56	96.6
Grade	29	100.0	56	96.6
Confirmation	27	93.1	45	77.6
Date of Dx	24	82.8	39	67.2*
Stage at Dx	6	20.7	1	1.7**
Date 1st Seen	25	86.2	51	87.9
Stage 1st Seen	5	17.2	1	1.7**

NOTE: Statistical significance of the tumor registrar performance versus the medical records personnel performance, using a one-degree-of-freedom chi-square test: * $p < 0.05$ ** $p < 0.01$.

TABLE 3: NUMBER AND PERCENT OF CORRECT RESPONSES BY TYPE OF RESPONDENT AND DATA ITEM FOR MEDICAL CHART 303030 - BREAST

CORRECT RESPONSES

DATA ITEM	TUMOR REGISTRARS		MEDICAL RECORDS PERSONNEL	
	NUMBER	PERCENT	NUMBER	PERCENT
Zip Code	29	100.0	58	100.0
County Residence	29	100.0	57	98.3
Date of Birth	29	100.0	57	98.3
Race	28	96.6	58	100.0
Ethnicity	22	78.6	21	36.2**
Sex	29	100.0	58	100.0
Soc. Sec. No.	29	100.0	58	100.0
Place of Birth	22	75.9	30	51.7*
Dx Elsewhere	26	89.7	49	87.5
Sequence #	24	82.8	14	24.1**
Primary Site	23	79.3	28	48.3**
PrimX (3 digits)	29	100.0	58	100.0
Laterality	29	100.0	57	98.3
Cell Type	26	89.7	55	94.8
Grade	29	100.0	57	98.3
Confirmation	22	75.9	51	87.9
Date of Dx	24	82.8	42	72.4
Stage at Dx	13	44.8	15	25.9
Date 1st Seen	28	96.6	53	91.4*
Stage 1st Seen	20	69.0	21	36.2**

NOTE: Statistical significance of the tumor registrar performance versus the medical records personnel performance, using a one-degree-of-freedom chi-square test: * $p < 0.05$ ** $p < 0.01$.

TABLE 4: NUMBER AND PERCENT OF CORRECT RESPONSES BY TYPE OF RESPONDENT AND DATA ITEM FOR MEDICAL CHART 404040 - LYMPHOMA

CORRECT RESPONSES

DATA ITEM	TUMOR REGISTRARS		MEDICAL RECORDS PERSONNEL	
	NUMBER	PERCENT	NUMBER	PERCENT
Zip Code	29	100.0	58	100.0
County of Residence	28	96.6	57	98.3
Date of Birth	29	100.0	56	96.6
Race	29	100.0	58	100.0
Ethnicity	22	78.6	22	37.9**
Sex	29	100.0	58	100.0
Soc. Sec. No.	27	93.1	58	100.0
Place of Birth	23	79.3	40	69.0
Dx Elsewhere	21	72.4	35	60.3
Sequence #	28	96.6	55	94.8
Primary Site	14	48.3	6	10.3**
PrimX (3 digits)	29	100.0	57	98.3
Laterality	24	82.8	35	60.3
Cell Type	22	75.9	13	22.4**
Grade	29	100.0	56	96.6
Confirmation	29	100.0	56	96.6
Date of Dx	27	93.1	44	75.9
Stage at Dx	11	37.9	7	12.1*
Date 1st Seen	21	72.4	26	44.8*
Stage 1st Seen	11	37.9	8	13.8*

NOTE: Statistical significance of the tumor registrar performance versus the medical records personnel performance, using a one-degree-of-freedom chi-square test: * $p < 0.05$ ** $p < 0.01$.

STATE LIBRARY OF NORTH CAROLINA
3 3091 00739 3754

TABLE 5: INCIDENCE RATES, MORTALITY RATES AND INCIDENCE/MORTALITY RATIOS FOR ALL CANCERS BY RACE-SEX GROUP, UNITED STATES, GEORGIA AND NORTH CAROLINA

RACE-SEX GROUP		UNITED STATES RATE	UNITED STATES RATIO	GEORGIA RATE	GEORGIA RATIO	NORTH CAROLINA RATE	NORTH CAROLINA RATIO
White Males	Incidence	433.1		430.6		438.7	
	Mortality	206.0	2.10	228.5	1.88	215.9	2.03
White Females	Incidence	339.8		320.7		360.7	
	Mortality	139.0	2.44	127.5	2.52	126.6	2.85
Nonwhite Males	Incidence	522.7		496.7		439.9	
	Mortality	301.6	1.73	298.3	1.66	303.9	1.45
Nonwhite Females	Incidence	324.9		296.0		293.9	
	Mortality	165.8	1.96	145.8	2.03	146.4	2.01
Total	Incidence	372.5		363.8		377.2	
	Mortality	166.8	2.23	174.7	2.08	170.6	2.21

NOTE: Georgia and United States data are 1987-1988 incidence rates and 1984-1988 mortality (SEER 1991). North Carolina data are 1990 incidence and 1984-1988 mortality. All rates are uniformly age-adjusted to the 1970 U.S. census. All rates are cases/100,000 and deaths/100,000 population.

Department of Environment, Health, and Natural Resources
State Center for Health and Environmental Statistics
P.O. Box 29538
Raleigh, N.C. 27626-0538
919/733-4728

BULK RATE
U.S. Postage
PAID
Raleigh, N.C. 27626-0538
Permit No. 1862

800 copies of this public document were printed at a cost of $120.96 or 0.15 per copy.